

**Ice Recrystallisation Inhibition Protein or Antifreeze Proteins**  
**from *Deschampsia*, *Lolium* and *Festuca* species of Grass**

- The present invention relates to nucleic acids or nucleic acid fragments encoding amino acid sequences for polypeptides involved in tolerance to freezing and/or low temperature stress in plants. More particularly, the present invention relates to nucleic acids or nucleic acid fragments encoding amino acid sequences for ice recrystallisation inhibition proteins (IRIPs) in plants, and the use thereof for the modification of plant response to freezing and/or low temperature stress. Even more particularly, the present invention relates to polypeptides involved in tolerance to freezing and/or low temperature stress in *Deschampsia* and *Lolium* species.
- 5
- 10 Plants have evolved a range of physiological and biochemical responses to freezing and low temperature stress. As a consequence of poikilothermy many plant species are tolerant of temperature extremes, including exposure to sub-zero temperatures. Sub-zero temperatures negatively impact on plant cells in many ways. As temperatures drop below freezing ice crystal formation initially takes place extracellularly, in the
- 15 apoplasm. This leads to an elevation of intracellular solute concentration as water is lost by osmosis to the extracellular ice, resulting in severe dehydration. Desiccation, whereby as much as 90% of intracellular water can be lost at  $-10^{\circ}\text{C}$ , induces multiple forms of membrane damage. Furthermore, extracellular ice obstructs gas and solute exchange, and growing ice crystals cause plasmolysis.
- 20 Plants and other organisms that are exposed to subzero temperatures have evolved varied mechanisms to confer tolerance to freezing stress including deployment of variant isozymes, synthesis of osmoprotectants and compatible solutes, and modification of membrane lipid composition. A particular characteristic of tolerance to freezing, and to temperature stresses in general, is the phenomenon of acquired
- 25 tolerance. For freezing stress this is termed cold acclimation, whereby a transition to low, non-freezing temperature can confer tolerance of subsequent exposure to otherwise lethal subzero temperatures.

A common response of plant and other species with tolerance to subzero temperatures and freezing is the expression of anti-freeze proteins (AFPs). AFPs have an affinity for

ice, by virtue of structural complementarity, thereby inhibiting its growth. Adsorption of AFPs onto ice surfaces has two distinct effects: thermal hysteresis (TH) and recrystallisation inhibition (RI). TH results from a noncolligative freezing point depression as ice front growth becomes restricted to sterically unfavourable spaces  
5 between AFPs. This broadens the gap between the melting and freezing points of ice, and this range is the measure of TH. AFPs mediate the effect of RI by interfering with the migration of ice boundaries which normally thermodynamically favour the creation of large, ice crystals at the expense of smaller ones. Thus RI activity limits the growth of large ice crystals that have the potential to puncture cell walls and membranes and  
10 cause plasmolysis. RI activity has been identified in extracts from a limited number of plant species, and the nucleotide sequence of one ice recrystallisation inhibition protein (IRIP) conferring such activity has been reported from *Lolium perenne*.

Antarctic hair grass *Deschampsia antarctica* is one of only two angiosperms to have overcome the geographical and environmental impediments to colonising the Antarctic  
15 continent. It grows in favourable locations along the western coast of the Antarctic Peninsula. *D. antarctica* is an over-wintering species with a short growing season that at Palmer Station (64°47'S), is typically November to March. In respect of low temperature stress, on Léonie Island in northern Marguerite Bay (67°36'S) towards the southern limit of distribution of *D. antarctica*, air temperatures below -30°C have been  
20 recorded during the austral winter. During the growing season, when plants are most vulnerable to freezing stress, episodes of temperatures down to -15°C can occur early in the growing season. *D. antarctica* has a well developed cold-acclimation response, and significant cellular damage only occurs in plants exposed to temperatures substantially below those at which they freeze.

25 Despite *D. antarctica*'s well developed freezing tolerance no biochemical or physiological mechanisms have been identified that can coherently account for this capacity.

There is a need for materials useful in modifying the tolerance to freezing and low temperature stress in a wide range of plants, and for methods for their use.

It is an object of the present invention to overcome, or at least alleviate, one or more of the difficulties or deficiencies associated with the prior art.

In one aspect, the present invention provides substantially purified or isolated nucleic acids or nucleic acid fragments encoding IRIPs from a *Deschampsia* species, preferably  
5 Antarctic hair-grass, *Deschampsia antarctica*, or functionally active fragments or variants thereof.

In a preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 8, 9, 11, 12, 14, 15, 17,  
10 18, 20, 21, 23 and 24 hereto; (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c); and (e) RNA sequences corresponding to the sequences recited in (a), (b), (c) and (d).

In another aspect, the present invention provides substantially purified or isolated  
15 nucleic acids or nucleic acid fragments encoding IRIPs from a ryegrass (*Lolium*) or fescue (*Festuca*) species. These species may be of any suitable type, including Italian or annual ryegrass, perennial ryegrass, tall fescue, meadow fescue and red fescue. Preferably the species is a ryegrass, more preferably perennial ryegrass (*L. perenne*).

In a preferred embodiment of this aspect of the invention, the substantially purified or  
20 isolated nucleic acid or nucleic acid fragment includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 26, 27, 29 and 30 hereto; (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c); and (e) RNA sequences corresponding to the  
25 sequences recited in (a), (b), (c) and (d).

The present invention provides substantially purified or isolated nucleic acids or nucleic acid fragments encoding amino acid sequences for a class of polypeptides which are related to IRIP or functionally active fragments or variants thereof. Such proteins are

referred to herein as IRIP-like. The genes which encode these polypeptides are expressed in a similar manner to IRIP. The invention also encompasses functionally active fragments and variants of nucleic acids encoding such polypeptides.

5 The individual or simultaneous enhancement or otherwise manipulation of IRIP or like gene activities in plants may enhance or otherwise alter the freezing and/or low temperature tolerance of plants.

10 The modification of plant freezing and/or low temperature tolerance based on the individual or simultaneous enhancement or otherwise manipulation of IRIP or like gene activities in plants has significant consequences for a range of applications in plant production and plant protection. For example, it has applications in increasing the range and productivity of plants.

Methods for the modification of plant freezing and/or low temperature tolerance may facilitate the production of, for example, plants with enhanced tolerance of freezing and/or low temperature stress.

15 Nucleic acids according to the invention may be full-length genes or part thereof, and are also referred to as "nucleic acid fragments" and "nucleotide sequences" in this specification.

20 The nucleic acid or nucleic acid fragment may be of any suitable type and includes DNA (such as cDNA or genomic DNA) and RNA (such as mRNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases, and combinations thereof.

Such nucleic acid fragments could be assembled to form a consensus contig.

25 In a further aspect of the present invention there is provided a substantially purified or isolated regulatory element from a *Deschampsia* species, preferably *Deschampsia antarctica*, said regulatory element being capable of modifying plant response to freezing and/or low temperature stress.

More preferably the regulatory element includes a nucleotide sequence selected from the groups consisting of (a) sequences shown in Figures 32 and 33 hereto; (b) complements of the sequences recited in (a); and (c) functionally active fragments and variants of the sequences recited in (a) and (b).

- 5 In a further aspect of the present invention there is provided a substantially purified or isolated regulatory element from a *Lolium* or *Festuca* species, preferably *Lolium perenne*, said regulatory element being capable of modifying plant response to freezing and/or low temperature stress.

- 10 More preferably the regulatory element includes a nucleotide sequence selected from the group consisting of (a) sequence shown in Figure 34 hereto; (b) complement of the sequence recited in (a) and (c) functionally active fragments and variants of the sequences recited in (a) and (b).

Preferably the regulatory element is a promoter.

- 15 Preferably the regulatory element is isolated from an IRIP nucleic acid or nucleic acid fragment.

As used herein, the term IRIP-like relates to polypeptides that are produced in the plant in substantially the same organs and at substantially the same developmental stages as IRIP.

- 20 The term "isolated" means that the material is removed from its original environment (eg. the natural environment if it is naturally occurring). For example, a naturally occurring nucleic acid or polypeptide present in a living plant is not isolated, but the same nucleic acid or polypeptide separated from some or all of the coexisting materials in the natural system, is isolated. Such nucleic acids could be part of a vector and/or such nucleic acids could be part of a composition, and still be isolated in that such a  
25 vector or composition is not part of its natural environment.

As used herein, the term "consensus contig" refers to a nucleotide sequence that is assembled from two or more constituent nucleotide sequences that share common or overlapping regions of sequence homology. For example, the nucleotide sequence of two or more nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences (and thus their corresponding nucleic acid fragments) can be assembled into a single contiguous nucleotide sequence.

The term "purified" means that the nucleic acid or polypeptide is substantially free of other nucleic acids or polypeptides.

By "functionally active" in respect of a nucleic acid it is meant that the fragment or variant (such as an analogue, derivative or mutant) is capable of modifying the tolerance of freezing and/or low temperature stress in a plant. Such variants include naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 80% identity to the relevant part of the above mentioned sequence, more preferably at least approximately 90% identity, most preferably at least approximately 95% identity. Such functionally active variants and fragments include, for example, those having nucleic acid changes which result in conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least 30 nucleotides, more preferably at least 45 nucleotides, most preferably at least 60 nucleotides.

By "functionally active" in respect of a polypeptide is meant that the fragment or variant has one or more of the biological properties of an IRIP or IRIP-like protein. Additions, deletions, substitutions and derivatizations of one or more of the amino acids are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 60% identity to the relevant part of the above mentioned sequence,

more preferably at least approximately 80% identity, most preferably at least approximately 90% identity. Such functionally active variants and fragments include, for example, those having conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least  
5 10 amino acids, more preferably at least 15 amino acids, most preferably at least 20 amino acids.

The term "construct" as used herein refers to an artificially assembled or isolated nucleic acid molecule which includes the gene of interest. In general a construct may include the gene or genes of interest, a marker gene which in some cases can also be the gene  
10 of interest and appropriate regulatory sequences. It should be appreciated that the inclusion of regulatory sequences in a construct is optional, for example, such sequences may not be required in situations where the regulatory sequences of a host cell are to be used. The term construct includes vectors but should not be seen as being limited thereto.

15 The term "vector" as used herein encompasses both cloning and expression vectors. Vectors are often recombinant molecules containing nucleic acid molecules from several sources.

By "operatively linked" is meant that said regulatory element is capable of causing expression of said nucleic acid or nucleic acid fragment in a plant cell and said  
20 terminator is capable of terminating expression of said nucleic acid or nucleic acid fragment in a plant cell. Preferably, said regulatory element is upstream of said nucleic acid or nucleic acid fragment and said terminator is downstream of said nucleic acid or nucleic acid fragment.

By "an effective amount" it is meant an amount sufficient to result in an identifiable  
25 phenotypic trait in said plant, or a plant, plant seed or other plant part derived therefrom. Such amounts can be readily determined by an appropriately skilled person, taking into account the type of plant, the route of administration and other relevant factors. Such a person will readily be able to determine a suitable amount and method of administration. See, for example, Maniatis *et al*, Molecular Cloning: A Laboratory Manual, Cold Spring

Harbor Laboratory, Cold Spring Harbor, the relevant disclosure of which is incorporated herein by reference.

It will also be understood that the term "comprises" (or its grammatical variants) as used in this specification is equivalent to the term "includes" and should not be taken as  
5 excluding the presence of other elements or features.

Genes encoding other IRIP or IRIP-like proteins for modifying the tolerance of plants to freezing and/or low temperature stress, either as cDNAs or genomic DNAs, may be isolated directly by using all or a portion of the nucleic acids or nucleic acid fragments of the present invention as hybridisation probes to screen libraries from the desired plant  
10 employing the methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the nucleic acid sequences of the present invention may be designed and synthesized by methods known in the art. Moreover, the entire sequences may be used directly to synthesize DNA probes by methods known to the skilled artisan, such as random primer DNA labelling, nick translation, or end-labelling  
15 techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers may be designed and used to amplify a part or all of the sequences of the present invention. The resulting amplification products may be labelled directly during amplification reactions or labelled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate  
20 stringency.

In addition, short segments of the nucleic acids or nucleic acid fragments of the present invention may be used in protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. For example, polymerase chain reaction may be performed on a library of cloned nucleic acid fragments wherein the sequence of one  
25 primer is derived from the nucleic acid sequences of the present invention, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, those skilled in the art can follow the RACE protocol [Frohman *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:8998, the entire disclosure of which is  
30



incorporated herein by reference] to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Using commercially available 3' RACE and 5' RACE systems (BRL), specific 3' or 5' cDNA fragments may be isolated [Ohara *et al.* (1989) *Proc. Natl. Acad Sci USA* 86:5673; Loh  
5 *et al.* (1989) *Science* 243:217, the entire disclosures of which are incorporated herein by reference]. Products generated by the 3' and 5' RACE procedures may be combined to generate full-length cDNAs.

In a further aspect of the present invention there is provided a substantially purified or isolated IRIP or IRIP-like polypeptide from a *Deschampsia* species, preferably from  
10 Antarctic hair-grass, *Deschampsia antarctica*; and functionally active fragments and variants thereof.

In a preferred embodiment of this aspect of the invention, the substantially purified or isolated polypeptide includes an amino acid sequence selected from the group consisting of sequences shown in Figures 10, 13, 16, 19, 22 and 25 hereto; and  
15 functionally active fragments and variants thereof.

In a still further aspect of the present invention there is provided a substantially purified or isolated IRIP or IRIP-like polypeptide from a ryegrass (*Lolium*) or fescue (*Festuca*) species; and functionally active fragments and variants thereof.

The ryegrass (*Lolium*) or fescue (*Festuca*) species may be of any suitable type,  
20 including Italian or annual ryegrass, perennial ryegrass, tall fescue, meadow fescue and red fescue. Preferably the species is a ryegrass, more preferably perennial ryegrass (*L. perenne*).

In a preferred embodiment of this aspect of the invention, the substantially purified or isolated polypeptide includes an amino acid sequence selected from the group  
25 consisting of sequences shown in Figures 28 and 31 hereto; and functionally active fragments and variants thereof.

The Applicant has found that the polypeptides of the present invention include relatively few leucine rich repeat (LRR) motifs.

Preferably said LLR motifs from a *Deschampsia* species include the consensus sequence:

5 LxLxxNxLTGxIPxxLGxLxxLxx (SEQ ID NO. 128)

or the consensus sequence:

LxLxxNxLSGxIPxxLGxLxxLxx (SEQ ID NO. 143)

Preferably said LRR motifs from a *Lolium* or *Festuca* species include the consensus sequence:

10 LxLxxNxLTGxIPxxLGxLxxLxx (SEQ ID NO. 129)

or the consensus sequence:

LxLxxNxLSGxIPxxLGxLxxLxx (SEQ ID NO. 144)

Applicant has found that polypeptides of the present invention including relatively few LRR motifs, preferably 3 or fewer LRR motifs, more preferably 1 or fewer LRR motifs, may be more efficient at modifying tolerance of freezing and/or low temperature stress in a plant than nucleic acids or nucleic acid fragments having relatively more LRR motifs, for example approximately 9 or more LRR motifs. Similarly, the nucleic acids or nucleic acid fragments encoding such polypeptides may be more efficient at modifying tolerance of freezing and/or low temperature stress in a plant.

20 In a further embodiment of this aspect of the invention, there is provided a polypeptide recombinantly produced from a nucleic acid or nucleic acid fragment according to the present invention. Techniques for recombinantly producing polypeptides are known to those skilled in the art.

Availability of the nucleotide sequences of the present invention and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides may be used to immunise animals to produce polyclonal or  
5 monoclonal antibodies with specificity for peptides and/or proteins comprising the amino acid sequences. These antibodies may be then used to screen cDNA expression libraries to isolate full-length cDNA clones of interest.

A genotype is the genetic constitution of an individual or group. Variations in genotype are essential in commercial breeding programs, in determining parentage, in  
10 diagnostics and fingerprinting, and the like. Genotypes can be readily described in terms of genetic markers. A genetic marker identifies a specific region or locus in the genome. The more genetic markers, the finer defined is the genotype. A genetic marker becomes particularly useful when it is allelic between organisms because it then may serve to unambiguously identify an individual. Furthermore, a genetic marker becomes  
15 particularly useful when it is based on nucleic acid sequence information that can unambiguously establish a genotype of an individual and when the function encoded by such nucleic acid is known and is associated with a specific trait. Such nucleic acids and/or nucleotide sequence information including single nucleotide polymorphisms (SNP's), variations in single nucleotides between allelic forms of such nucleotide  
20 sequence, can be used as perfect markers or candidate genes for the given trait.

Applicants have identified a number of SNPs of the nucleic acids or nucleic acid fragments of the present invention. These are indicated (marked with grey on the black background) in the figures that show multiple alignments of nucleotide sequences of nucleic acid fragments contributing to consensus contig sequences. See for example,  
25 Figures 8, 11, 14, 17, 20, 23, 26 and 29 hereto.

Accordingly, in a further aspect of the present invention, there is provided a substantially purified or isolated nucleic acid or nucleic acid fragment including a single nucleotide polymorphism (SNP) from a nucleic acid fragment shown in Figures 8, 9, 11, 12, 14, 15, 17, 18, 20, 21, 23, 24, 26, 27, 29 and 30 hereto, or complements or sequences  
30 antisense thereto.

In a still further aspect of the present invention there is provided a method of isolating a nucleic acid or nucleic acid fragment of the present invention including a single nucleotide polymorphism (SNP), said method including sequencing nucleic acid fragments from a nucleic acid library.

- 5 The nucleic acid library may be of any suitable type and is preferably a cDNA library.

The nucleic acid fragments may be isolated from recombinant plasmids or may be amplified, for example using polymerase chain reaction.

The sequencing may be performed by techniques known to those skilled in the art.

- 10 In a still further aspect of the present invention, there is provided use of nucleic acids or nucleic acid fragments of the present invention including SNP's, and/or nucleotide sequence information thereof, as molecular genetic markers.

In a still further aspect of the present invention there is provided use of a nucleic acid or nucleic acid fragment according to the present invention, and/or nucleotide sequence information thereof, as a molecular genetic marker.

- 15 More particularly, nucleic acids or nucleic acid fragments according to the present invention and/or nucleotide sequence information thereof may be used as a molecular genetic marker for quantitative trait loci (QTL) tagging, QTL mapping, DNA fingerprinting and in marker assisted selection, particularly in grasses and cereals. Even more particularly, nucleic acids or nucleic acid fragments according to the present invention  
20 and/or nucleotide sequence information thereof may be used as molecular genetic markers in grass and cereal improvement, e.g. tagging QTLs for tolerance to freezing and/or low temperature stress. Even more particularly, sequence information revealing SNPs in allelic variants of the nucleic acids or nucleic acid fragments of the present invention and/or nucleotide sequence information thereof may be used as molecular  
25 genetic markers for QTL tagging and mapping and in marker assisted selection, particularly in grasses and cereals.

In a still further aspect of the present invention there is provided a construct including a nucleic acid or nucleic acid fragment according to the present invention. The construct may be a vector.

5 In a preferred embodiment of this aspect of the invention, the vector may include a regulatory element such as a promoter, a nucleic acid or nucleic acid fragment according to the present invention and a terminator; said regulatory element, nucleic acid or nucleic acid fragment and terminator being operatively linked.

10 In a further preferred embodiment of this aspect of the invention, the vector may include a regulatory element according to the present invention, a further nucleic acid molecule and a terminator; said regulatory element, further nucleic acid molecule and terminator being operatively linked.

15 In a still further preferred embodiment of this aspect of the invention, the vector may include a regulatory element according to the present invention, a nucleic acid or nucleic acid fragment according to the present invention and a terminator, said regulatory element, nucleic acid or nucleic acid fragment and terminator being operatively linked.

20 The vector may be of any suitable type and may be viral or non-viral. The vector may be an expression vector. Such vectors include chromosomal, non-chromosomal and synthetic nucleic acid sequences, eg. derivatives of plant viruses; bacterial plasmids; derivatives of the Ti plasmid from *Agrobacterium tumefaciens*, derivatives of the Ri plasmid from *Agrobacterium rhizogenes*; phage DNA; yeast artificial chromosomes; bacterial artificial chromosomes; binary bacterial artificial chromosomes; vectors derived from combinations of plasmids and phage DNA. However, any other vector may be used as long as it is replicable, or integrative or viable in the plant cell.

25 The regulatory element and terminator may be of any suitable type and may be endogenous to the target plant cell or may be exogenous, provided that they are functional in the target plant cell.

In another embodiment, the construct or vector may include more than one nucleic acid. The nucleic acids within the same construct or vector may have identical or differing sequences. In one preferred embodiment, the construct or vector has at least two nucleic acids encoding functionally similar enzymes. In a particularly preferred  
5 embodiment, each further nucleic acid molecule has one or more upstream regulatory elements and one or more downstream terminators, although expression of more than one further nucleic acid molecule from an upstream regulatory element or termination of more than one further nucleic acid molecule from a downstream terminator(s) is not precluded.

10 Preferably the regulatory element is a promoter. A variety of promoters which may be employed in the constructs and vectors of the present invention are well known to those skilled in the art. Factors influencing the choice of promoter include the desired tissue specificity of the vector, and whether constitutive or inducible expression is desired and the nature of the plant cell to be transformed (eg. monocotyledon or dicotyledon).

15 Particularly suitable promoters include but are not limited to the constitutive Cauliflower Mosaic Virus 35S (CaMV 35S) promoter and derivatives thereof, the maize *Ubiquitin* promoter, the rice *Actin* promoter, and the tissue-specific *Arabidopsis* small subunit (ASSU) promoter. Alternatively, the regulatory element may be a regulatory element according to the present invention.

20 A variety of terminators which may be employed in the vectors and constructs of the present invention are also well known to those skilled in the art. The terminator may be from the same gene as the promoter sequence or a different gene. Particularly suitable terminators are polyadenylation signals, such as the CaMV 35S polyA and other terminators from the *nopaline synthase* (*nos*), the *octopine synthase* (*ocs*) and the *rbcS*  
25 genes.

The further nucleic acid molecule may be a sequence, for example a gene or fragment thereof, or sequence antisense thereto, which is capable of modifying plant response to freezing and/or low temperature stress. It may be a nucleic acid or nucleic acid fragment according to the present invention, but is not limited thereto.

The vector, in addition to the regulatory element, the nucleic acid or nucleic acid fragment of the present invention and the terminator, may include further elements necessary for expression of the nucleic acid or nucleic acid fragment, in different combinations, for example vector backbone, origin of replication (*ori*), multiple cloning  
5 sites, recognition sites for recombination events, spacer sequences, enhancers, introns (such as the maize *Ubiquitin (Ubi)* intron), antibiotic resistance genes and other selectable marker genes [such as the *neomycin phosphotransferase (npt2)* gene, the *hygromycin phosphotransferase (hph)* gene, the *phosphinothricin acetyltransferase (bar* or *pat)* gene and the *gentamycin acetyl transferase (aacC1)* gene], and reporter genes  
10 (such as *beta-glucuronidase (GUS)* gene (*gusA*) and *green fluorescent protein (gfp)*]. The vector may also contain a ribosome binding site for translation initiation. The vector may also include appropriate sequences for amplifying expression.

As an alternative to use of a selectable marker gene to provide a phenotypic trait for selection of transformed host cells, the presence of the construct or vector in  
15 transformed cells may be determined by other techniques well known in the art, such as PCR (polymerase chain reaction), Southern blot hybridisation analysis, histochemical GUS assays, visual examination including microscopic examination of fluorescence emitted by *gfp*, northern and Western blot hybridisation analyses.

Those skilled in the art will appreciate that the various components of the construct or  
20 vector are operatively linked, so as to result in expression of said nucleic acid or nucleic acid fragment. Techniques for operatively linking the components of the vector of the present invention are well known to those skilled in the art. Such techniques include the use of linkers, such as synthetic linkers, for example including one or more restriction enzyme sites.

25 The constructs and vectors of the present invention may be incorporated into a variety of plants, including monocotyledons (such as grasses from the genera *Deschampsia*, *Lolium*, *Festuca*, *Paspalum*, *Pennisetum*, *Panicum* and other forage and turf grasses, corn, oat, sugarcane, wheat and barley), dicotyledons (such as *Arabidopsis*, tobacco, white clover, red clover, subterranean clover, alfalfa, eucalyptus, potato, sugarbeet,  
30 canola, soybean, chickpea) and gymnosperms.

Techniques for incorporating the constructs and vectors of the present invention into plant cells (for example by transduction, transfection or transformation) are well known to those skilled in the art. Such techniques include *Agrobacterium*-mediated introduction, electroporation to tissues, cells and protoplasts, protoplast fusion, injection  
5 into reproductive organs, injection into immature embryos and high velocity projectile introduction to cells, tissues, calli, immature and mature embryos. The choice of technique will depend largely on the type of plant to be transformed.

Cells incorporating the constructs and vectors of the present invention may be selected, as described above, and then cultured in an appropriate medium to regenerate  
10 transformed plants, using techniques well known in the art. The culture conditions, such as temperature, pH and the like, will be apparent to the person skilled in the art. The resulting plants may be reproduced, either sexually or asexually, using methods well known in the art, to produce successive generations of transformed plants.

In a further aspect of the present invention there is provided a plant cell, plant, plant  
15 seed or other plant part, including, e.g. transformed with, a construct or vector of the present invention.

The plant cell, plant, plant seed or other plant part may be from any suitable species, including monocotyledons, dicotyledons and gymnosperms.

The present invention also provides a plant, plant seed or other plant part, or a plant  
20 extract, derived from a plant cell of the present invention.

The present invention also provides a plant, plant seed or other plant part, or a plant extract, derived from a plant of the present invention.

In a further aspect of the present invention there is provided a method of modifying tolerance of freezing and/or low temperature stress in a plant, said method including  
25 introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment, construct and/or a vector according to the present invention.



Using the methods and materials of the present invention, the tolerance of freezing and/or low temperature stress in a plant may be increased or decreased or otherwise modified. For example, the tolerance of freezing and/or low temperature stress may be increased or otherwise altered. They may be increased, for example, by incorporating  
5 additional copies of a sense nucleic acid or nucleic acid fragment of the present invention. They may be decreased, for example, by incorporating an antisense nucleic acid or nucleic acid fragment of the present invention.

In a further aspect of the present invention there is provided a preparation for transforming a plant comprising at least one nucleic acid according to the present  
10 invention. The preparation may contain vectors or other constructs to facilitate administration to and/or transformation of the plant with the nucleic acid.

The present invention will now be more fully described with reference to the accompanying Examples and drawings. It should be understood, however, that the description following is illustrative only and should not be taken in any way as a  
15 restriction on the generality of the invention described above.

#### In the Figures

Figure 1. RI assay on total extracts of leaves from non-acclimated (grown at 22°C) and cold acclimated (5°C) *D. antarctica*. A, Initial ice crystal structure following snap freezing. B, Ice crystal structure after 16h incubation at -3°C.  
20 Capillary B contains extraction buffer; capillaries 1-7: 1000, 250, 62.5, 15.6, 3.91, 0.977 and 0.244  $\mu\text{g mL}^{-1}$  respectively of total leaf protein. Extracts were either untreated or subject to incubation at 95°C for 5 min as indicated. Endpoint of RI activity defined as the lowest protein concentration ( $\mu\text{g mL}^{-1}$ ) at which ice crystal structure in B remains unchanged from that in A.

25 Figure 2. A, Repeat structures of representative IRIP orthologues HvIRIP (SEQ ID NO. 126) and DalRIPd (SEQ ID NO. 31). Four cysteine residues conserved in LRR proteins, and predicted to form two disulphide bridges are shown in bold, connected by lines to show predicted bridges. A highly conserved 9 amino acid

motif including 3 of these cysteine residues is underlined. Consensus sequences for plant LRRs (SEQ ID NOs. 128 and 143) (Kobe and Kajava (2001) Curr Opin Struct Biol 11:725), and the IRIP repeat (SEQ ID NO. 141) (this work) are shown in bold below the tandem repeat alignments, and identical residues highlighted by grey boxes. B, Sequence alignment of IRIP orthologues and a putative PSKR orthologue from *Oryza sativa*. Sequences include *LpIRIP* (Sidebottom, et al. (2000 Nature 406:256)) (SEQ ID NO. 124), predicted IRIP orthologue *TalIRIP* derived from assembly of *T. aestivum* sequences in the NCBI EST database (SEQ ID NO. 125), predicted *HvIRIP* derived from assembly of *H. vulgare* sequences in the NCBI EST database (SEQ ID NO. 126), predicted *LmlIRIP* derived from assembly of *L. multiflorum* sequences in the NCBI EST database (SEQ ID NO. 142), and OsPSKR a putative PSKR orthologue from *Oryza sativa* (NP\_911036) (SEQ ID NO. 127). Sequences of the present invention are *DalRIPa* (SEQ ID NO. 17), *DalRIPd* (SEQ ID NO. 31) *DalRIPe.7* (SEQ ID NO. 38), *LpIRIPa* (SEQ ID NO. 102) and *LpIRIPb* (SEQ ID NO. 120). Identical and conserved residues are highlighted by black and grey boxes respectively. Four invariantly conserved cysteine residues are marked with arrowheads. C, Schematic of domain organisation in IRIP orthologues and OsPSKR. SP: signal peptide; 2x S-S; domain predicted to form 2 disulphide bridges; LRRs numbered; Island/Isa./Is.: island domain; IRIP repeats unlabelled; TM: transmembrane domain. D, Phylogenetic tree of IRIP orthologues and LRR proteins. Branch lengths are proportional to the number of amino acid substitutions per site (indicated by scale bar). LRR proteins include accession number as suffix.

Figure 3. Structural modelling of IRIPs. A, Theoretical structure of *DalRIPa* aligned along the prism face of ice (parallel to the *a*-axis). B, Theoretical structure of *LpIRIPa* aligned along the prism face of ice (parallel to the *a*-axis). Cysteine residues at positions 120 and 143 have been modelled to participate in an additional disulphide bond, relative to *DalRIPa*. C, Ribbon backbone diagram of *DalRIPa* highlighting the amino acid residue composition of the two putative ice binding surfaces, side "A" and side "B". D, Cross-sectional view of ribbon

backbone of 2  $\beta$ -roll loops of *DaIRIPa* showing positions of amino acid residue side chains.

Figure 4. Genomic organisation of *IRIP* genes. A, *D. antarctica* genomic Southern probed with *DaIRIPe*. B, *L. perenne* polygenic genomic Southern probed with *LpIRIPa*. C, *L. perenne* isogenic genomic Southern probed with *LpIRIPa*. 1: Undigested; 2: *SphI*; 3: *KpnI*; 4: *HindIII*; 5: *BamHI*.

Figure 5. SNP-based genetic mapping. A, Distribution of locus-specific amplification primers, genomic amplicons and putative SNP loci across the components of the *LpIRIPa* transcriptional unit. *LpIRIP* SNPs are indicated using the nomenclature na (number), indicating that the SNP was identified between NA6 parental haplotypes, and the relevant base pair coordinate. SNPs derived from single gene copies that segregate in the  $F_1$  progeny are shown in bold, while SNPs that potentially discriminate between paralogous gene copies are shown in normal text. B, Genetic map order in the upper part of the LG1 of the NA<sub>6</sub> parental genetic map, showing the *LpIRIPNA476*-detected SNP locus.

Figure 6. Analysis of *IRIP* gene expression in response to temperature. A, Northern blot of RNA from *D. antarctica* leaves and roots grown at 22, 4, and -16°C, probed with *DaIRIPe*. B, Northern blot of RNA from *L. perenne* leaves and roots grown at 22 and 4°C, probed with *LpIRIPa*.

Figure 7. RI assay on total extracts of *E. coli* expressing a putative orthologue of histone H3.2 and *DaIRIPe*. A, Initial ice crystal structure following snap freezing. B, Ice crystal structure after 16h incubation at -3°C. Capillary EB contains extraction buffer; capillary BSA 1000  $\mu\text{g mL}^{-1}$  bovine serum albumin; capillary PC cold acclimated *D. antarctica* leaf extract as positive control; capillaries 1-4: 400, 100, 25 and 6.25  $\mu\text{g mL}^{-1}$  respectively of total extracts of *E. coli*.

Figure 8. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *DaIRIPa* (SEQ ID NOs. 1-15)

Figure 9. Consensus nucleotide sequence of *DalRIPa* (SEQ ID NO. 16)

Figure 10. Deduced amino acid sequence of *DalRIPa* (SEQ ID NO. 17)

Figure 11. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *DalRIPb* (SEQ ID NOs. 18-20)

5 Figure 12. Consensus nucleotide sequence of *DalRIPb* (SEQ ID NO. 21)

Figure 13. Deduced amino acid sequence of *DalRIPb* (SEQ ID NO. 22)

Figure 14. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *DalRIPd* (SEQ ID NOs. 23-29)

Figure 15. Consensus nucleotide sequence of *DalRIPd* (SEQ ID NO. 30)

10 Figure 16. Deduced amino acid sequence of *DalRIPd* (SEQ ID NO. 31)

Figure 17. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *DalRIPe7* (SEQ ID NOs. 32-36)

Figure 18. Consensus nucleotide sequence of *DalRIPe7* (SEQ ID NO. 37)

Figure 19. Deduced amino acid sequence of *DalRIPe7* (SEQ ID NO. 38)

15 Figure 20. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *DalRIPe8* (SEQ ID NOs. 39-44)

Figure 21. Nucleotide sequence of *DalRIPe8* (SEQ ID NO. 45)

Figure 22. Deduced amino acid sequence of *DalRIPe8* (SEQ ID NO. 46)

20 Figure 23. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *DalRIPf* (SEQ ID NOs. 47-52)

Figure 24. Consensus nucleotide sequence of *DalRIPf* (SEQ ID NO. 53)

Figure 25. Deduced amino acid sequence of *DalRIPf* (SEQ ID NO. 54)

Figure 26. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *LpIRIPa* (SEQ ID NOs. 55-100)

5 Figure 27. Consensus nucleotide sequence of *LpIRIPa* (SEQ ID NO. 101)

Figure 28. Deduced amino acid sequence of *LpIRIPa* (SEQ ID NO. 102)

Figure 29. Nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence *LpIRIPb* (SEQ ID NOs. 103-118)

Figure 30. Consensus nucleotide sequence of *LpIRIPb* (SEQ ID NO. 119)

10 Figure 31. Deduced amino acid sequence of *LpIRIPb* (SEQ ID NO. 120)

Figure 32. Nucleotide sequence of promoter region of *DalRIPa* extending to the initiating ATG (underlined) (SEQ ID NO. 121)

Figure 33. Nucleotide sequence of promoter region of *DalRIPd* extending to the initiating ATG (underlined) (SEQ ID NO. 122)

15 Figure 34. Nucleotide sequence of promoter region of *LpIRIPa* extending to the initiating ATG (underlined) (SEQ ID NO. 123)

Figure 35. Plasmid map of vector used for *DalRIPa* gain of function biolistic transformation of wheat.

20 Figure 36. Plasmid map of vector used for *DalRIPd* gain of function biolistic transformation of wheat.

Figure 37. Plasmid map of vector used for *DalRIPe7* gain of function biolistic transformation of wheat.

Figure 38. Plasmid map of vector used for *DalRIPe8* gain of function biolistic transformation of wheat.

5 Figure 39. Plasmid map of vector used for *LpIRIPa* gain of function biolistic transformation of ryegrass.

Figure 40. Plasmid map of vector used for *LpIRIPb* gain of function biolistic transformation of ryegrass.

10 Figure 41. Plasmid map of vector used for *LpIRIPa* loss of function biolistic transformation of ryegrass.

Figure 42. Plasmid map of vector used for *LpIRIPb* loss of function biolistic transformation of ryegrass.

Figure 43. Plasmid map of vector used for *DalRIPa* gain of function *Agrobacterium* mediated transformation of wheat and barley.

15 Figure 44. Plasmid map of vector used for *DalRIPd* gain of function *Agrobacterium* mediated transformation of wheat and barley.

Figure 45. Plasmid map of vector used for *DalRIPe7* gain of function *Agrobacterium* mediated transformation of wheat and barley.

20 Figure 46. Plasmid map of vector used for *DalRIPe8* gain of function *Agrobacterium* mediated transformation of wheat and barley.

Figure 47. Plasmid map of vector used for *DalRIPa* gain of function *Agrobacterium* mediated transformation of *Arabidopsis*.

Figure 48. Plasmid map of vector used for *DalRIPd* gain of function *Agrobacterium* mediated transformation of *Arabidopsis*.

Figure 49. Plasmid map of vector used for *DalRIPe7* gain of function *Agrobacterium* mediated transformation of *Arabidopsis*.

5 Figure 50. Plasmid map of vector used for *DalRIPe8* gain of function *Agrobacterium* mediated transformation of *Arabidopsis*.

Figure 51. Plasmid map of vector used for *LpIRIPa* gain of function *Agrobacterium* mediated transformation of *Arabidopsis*.

10 Figure 52. Plasmid map of vector used for *LpIRIPb* gain of function *Agrobacterium* mediated transformation of *Arabidopsis*.

Figure 53. Plasmid map of vector used for *DalRIPd* promoter::GUS reporter gene *Agrobacterium* mediated transformation of *Arabidopsis*.

### Example 1

#### 15 Materials and Methods

##### Plant propagation, stress treatments, extraction of RI activity, DNA and RNA

*D. antarctica* material was collected in the vicinity of Jubany station on King George Island (62°14'S 58°40'W). Plants were germinated from seeds in the soil seed bank and thereafter were propagated vegetatively. *Lolium perenne* plants were of cultivar  
20 Impact. Doubled haploid *L. perenne* plants, where isolate DH297 of cultivar Verna.

Individual plants were grown in soil at the indicated temperatures under a 16/8h light/dark regime and photosynthetic photon flux intensity of 400  $\mu\text{mol m}^{-2} \text{s}^{-1}$  in

Enconair (Winnipeg, Canada) GC-20 plant growth chambers. Plants were cold acclimated by growth at 5°C for 2 weeks. Plants were divided into aerial and subterranean parts and snap frozen in liquid N<sub>2</sub>.

5 Total cellular extracts were prepared after (Doucet et al (2000) Cryobiology 40:218) by grinding plant material under liquid N<sub>2</sub> and resuspending the powder in 2 mL g<sup>-1</sup> of freshly prepared extraction buffer (50mM Tris pH 7.4, 20mM ascorbate, 10mM EDTA). The extract was filtered through miracloth (Calbiochem, La Jolla, USA). Apoplastic extracts were based on the method of Chun et al (1998) Euphytica 102:219. Leaf material was vacuum infiltrated in extraction buffer for 30 min, excess liquid removed  
10 from the leaves, and extracts collected by centrifugation at 800g for 30 min. All extracts were aliquoted, frozen in liquid N<sub>2</sub>, and stored at -80°C.

RNA and DNA were extracted using RNeasy and DNeasy Plant Mini kits (QIAGEN, Germany) respectively.

#### Ice recrystallisation inhibition (RI) assays

15 Extracts were thawed and insoluble material pelleted by centrifugation at 16,060g for 5 min. Supernatants were collected and for total cellular extracts protein content quantified using the Bio-Rad protein assay (Bio-Rad, Mississauga, ON, Canada), according to the manufacturer's instructions. Unless otherwise stated all extracts were incubated at 95°C for 5 min. The supernatants from both heat treated and untreated  
20 extracts were collected following centrifugation at 16,060g for 2 min. Serial 2 or 4-fold dilutions into extraction buffer were prepared.

The capillary method for the RI assay was modified from that of Tomczak et al (2003) Biochem. Biophys. Res. Commun. 311:1041. Briefly, extracts were loaded into 10µl glass capillaries (Drummond Scientific, Broomall, PA, USA), heat sealed and arrayed on  
25 a glass slide secured with adhesive tape. Extraction buffer, and BSA (Bio-Rad, Mississauga, ON, Canada) dissolved in extraction buffer to a final concentration of 1000 µg mL<sup>-1</sup> were included as negative controls. The capillary array was snap-frozen in an ethanol/dry ice bath and immersed in a reservoir of motor vehicle coolant diluted to a



final concentration of 10% monoethylene glycol, atop a jacketed stage through which the same solution at  $-3^{\circ}\text{C}$  was circulated using a refrigerated water bath (PolyScience Model 910, Niles, IL, USA).

- Samples were scored after overnight ( $\sim 16$  h) incubation at  $-3^{\circ}\text{C}$ . The endpoint of RI activity in total leaf extracts is defined as the lowest protein concentration ( $\mu\text{g mL}^{-1}$ ) at which ice crystal structure following incubation at  $-3^{\circ}\text{C}$  for 16h, remained unchanged from that initially induced by snap freezing. For apoplastic extracts, because of the low yields of protein from non-acclimated plants, the endpoint of RI activity was expressed as the equivalent wet weight of starting plant material per volume of extract.
- 10 Digital images were captured with a Leica DFC 300 F camera mounted on a Leica MZFL III stereoscopic zoom microscope using Leica FireCam software (Leica, Heerbrugg, Switzerland). Polarising light filters mounted perpendicularly on the microscope objective lens and beneath the stage enhanced visualisation of ice crystal morphology.
- 15 **Preparation of cDNA libraries, isolation and sequencing of cDNAs coding for IRIPs from Antarctic hair-grass, *Deschampsia antarctica*.**

cDNA libraries representing mRNAs from various organs and tissues from Antarctic hair-grass, *Deschampsia antarctica* were prepared. The characteristics of the libraries are described below (Table 1).

20

TABLE 1

cDNA libraries from Antarctic hair-grass, *Deschampsia antarctica*.

Library	Organ/Tissue
05Da	Aerial parts grown at $4^{\circ}\text{C}$
08Da	Roots grown at $-15^{\circ}\text{C}$
09Da	Roots transferred from $-15^{\circ}\text{C}$ to $25^{\circ}\text{C}$ for 24 h
10Da	Aerial parts transferred from $-15^{\circ}\text{C}$ to $25^{\circ}\text{C}$ for 24 h
11Da	Aerial parts grown at $-15^{\circ}\text{C}$

12Da	Roots grown at -15°C
15Da	Roots grown at 4°C
16Da	Aerial parts grown at 4°C
17Da	Roots transferred from 25°C to 0°C for 48 h
18Da	Aerial parts transferred from -15°C to 0°C for 48 h
19Da	Aerial parts transferred from 25°C to 0°C for 48 h, then to -15°C for 48 h
20Da	Aerial parts grown at -15°C
21Da	Aerial parts grown at 4°C
22Da	Roots grown at -15°C
23Da	Roots grown at 4°C

The cDNA libraries may be prepared by any of many methods available. For example, total RNA may be isolated using the Trizol method (Gibco-BRL, USA) or the RNeasy Plant Mini kit (Qiagen, Germany), following the manufacturers' instructions. cDNAs may be generated using the SMART PCR cDNA synthesis kit (Clontech, USA), cDNAs may be amplified by long distance polymerase chain reaction using the Advantage 2 PCR Enzyme system (Clontech, USA), cDNAs may be cleaned using the GeneClean spin column (Bio 101, USA), tailed and size fractionated, according to the protocol provided by Clontech. The cDNAs may be introduced into the pGEM-T Easy Vector system (Promega, USA) according to the protocol provided by Promega. The cDNAs in the pGEM-T Easy plasmid vector are transfected into *Escherichia coli* Epicurian coli XL10-Gold ultra competent cells (Stratagene, USA) according to the protocol provided by Stratagene.

Alternatively, the cDNAs may be introduced into plasmid vectors for first preparing the cDNA libraries in Uni-ZAP XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA, USA). The Uni-ZAP XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into pre-cut pBluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into *E. coli* DH10B cells according to the manufacturer's protocol (GIBCO BRL Products).

cDNA clones encoding putative IRIP orthologues from *D. antarctica* came from 5 libraries derived from either shoots or roots grown at 4°C or -15°C, and from shoots transferred from -15°C to 25°C for 24 hours. One of these variants (*DalRIPd*) was isolated from both root and shoot libraries, whilst the other four forms were derived from 5 shoot libraries only.

### Molecular cloning of genomic sequences

All but one of *D. antarctica* cDNAs encoding IRIP orthologues (*DalRIPa*), encoded N-terminally truncated hypothetical IRIP forms. Therefore full length genomic sequences were isolated using the GenomeWalker kit (BD Biosciences, Palo Alto, USA) and 10 nested gene specific primers. The 3' UTR primers for *DalRIPd* were:  
 5' GACATCGCGATTGGTCCCACCAAGTG 3' (SEQ ID NO. 130), and  
 5' GCATCCTGCACGGACATATCATT 3' (SEQ ID NO. 131);  
 and *DalRIPe*:  
 5' GTTACATAAGACGATTGGCCCCACCAAG 3' (SEQ ID NO. 132), and  
 15 5' CAATCCACTCACTGATCATTAAACCACC 3' (SEQ ID NO. 133). For the isolation of *LpIRIPa* nested primers  
 5' GATGCTATATCCACGAAGTTACAT 3' (SEQ ID NO. 134), and  
 5' ATTGGCCCCACCAAGTGA 3' (SEQ ID NO. 135) complementary to conserved regions within the 3' UTR of the *D. antarctica* IRIP forms were employed. *LpIRIPa* was 20 also obtained from the *L. perenne* North African<sub>6</sub> x Aurora<sub>6</sub> genetic mapping population (see below). PCR products were purified using QIAquick gel extraction kit (QIAGEN, Germany) and molecularly cloned into pGEM-T Easy as above.

### DNA template preparation, sequencing and analysis

25 Templates for sequencing of cDNA and genomic clones plasmid DNA was purified using a QIAprep spin miniprep kit (QIAGEN, Germany). Sequencing reactions primed with a modified SMART primer (5' AAGCAGTGGTAACAACGCAGAGTGGG 3') (SEQ ID NO. 136), M13F or M13R primers were performed either using BigDye Terminators or ET Terminators, and the reaction products resolved on an ABI Prism 3700, or 3730xl DNA Analyser (PE Biosystems, Foster City, USA), or a MegaBACE 4000 (Amersham

Biosciences, UK) respectively. Sequence files were used as queries for BLASTX, BLASTN and TBLASTN Altschul et al 1987 Nucleic Acids Res. 25:3389 searches of the SWISS-PROT, GenBank Main and GenBank EST databases respectively. BLASTX searches of the NCBI database of GenBank+EMBL+DDBJ sequences from EST divisions with full length *IRIP* gene sequences reveal the presence of many *IRIP* gene related sequences in grasses including *Lolium multiflorum* (Italian rye grass), *Leymus chinensis*, *Puccinellia tenuiflora*, and from in-house EST data *Agrostis tenuis*. We have also identified in the NCBI EST database orthologous sequences in the cereals *Hordeum vulgare* (barley), *S. cereale* (winter rye), and *Triticum monococcum*, *T. turgidum* and *T. aestivum* (diploid, tetraploid and hexaploid forms of wheat), frequently associated with cDNA libraries derived from cold-stressed or vernalised material. We are not yet aware of any species of the Pooideae without *IRIP* orthologues. EST sequences were also assembled to derive representative *IRIP* orthologues. Sequence assembly was performed using Sequencher (Gene Codes, Ann Arbor, MI, USA).

Potential signal sequences were identified by analysis with SignalP (<http://www.cbs.dtu.dk/services/SignalP/>). Subcellular localisation was predicted using TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) and PSORT (<http://psort.nibb.ac.jp>). Phylogenetic analysis was performed using Vector NTI (Invitrogen, Carlsbad, CA, USA) using the Align X algorithm with default parameters.

## 20 Structural modelling

Homology modelling was performed using Schrödinger molecular modelling software (version macromodel 8.6, Portland, OR, USA) using the Prime homology modelling module. A homology model of the *DalRIPa* and *LplRIPa* proteins was build using the N terminus of the crystal structure of *Phaseolus vulgaris* polygalacturonase-inhibiting protein (PGIP) (pdb entry: 1OGQ) and a right-handed version of a theoretical *Lolium perenne* AFP model (Kuiper et al (2001) *Biophys. J.* 81:3560) as templates. The model was geometrically optimised with distance constraints holding optimal hydrogen bond distances between beta sheet regions for 10,000 iterations using an OPLS2001 forcefield and Generalised Born (GB) solvation. This was followed by an additional 5000 iterations minimisation without constraints applied. Images were generated using Swiss Pdb Viewer and Pov-Ray3.5.

### Southern and Northern analysis

For Southern 10 µg of DNA was digested with restriction enzymes, separated on a 1% agarose TAE gel and stained with ethidium bromide before transfer to Hybond-N membrane (Amersham Biosciences). For Northern 10 µg of RNA was separated on a 1% agarose formaldehyde gel, transferred to Hybond-N membrane (Amersham Biosciences) and stained with methylene blue. Isolated fragments of *DaIRIPe7* and *LpIRIPa* were radio-labelled with  $\alpha$ -<sup>32</sup>P-dCTP using the Rediprime II Random Prime Labelling System (Amersham Biosciences) and purified with MicroSpin™ S-400 HR Columns (Amersham Biosciences) according to the manufacturer's instructions. The blotted membranes were hybridised with radio-labelled probe as described by Maniatis *et al*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor. Hybridisation patterns were visualised and quantified on a Typhoon 8600 Variable Mode Imager (Amersham Biosciences) according to the manufacturers instructions.

### 15 SNP-based genetic mapping

The perennial ryegrass population used for SNP discovery and genetic mapping was an F<sub>1</sub> progeny set derived from a reciprocal pair-cross between the heterozygous parental genotypes North African<sub>6</sub> (NA<sub>6</sub>) and Aurora<sub>6</sub> (AU<sub>6</sub>) (Faville *et al.*, *Theor Appl Genet.* in press).

20 The procedure for targeted *in vitro* SNP discovery in perennial ryegrass was described by Forster *et al.* (2004) Molecular marker-based genetic analysis of pasture and turf grasses, in: Molecular Breeding of Forage and Turf, Hopkins *et al* (eds) Kluwer Academic Press pp 197-239. The partial cDNA sequence of the *IRIP* reported from *L. perenne* (Sidebottom *et al* (2000) *Nature* 406:256), along with the sequence of *LpIRIPa*,  
25 were assembled into a single contig and locus amplification primers (LAPs) were designed to generate 3 amplicons located at various positions within the gene unit (Fig. 5A), to cover the 5'-untranslated region (UTR), single exon and 3'-UTR regions. Genomic amplicons were generated using standard PCR conditions from each parental genotype of the F<sub>1</sub>(NA<sub>6</sub> x AU<sub>6</sub>) population. PCR fragments were cloned using the

TOPO TA system (Invitrogen K4575-40) and DNA sequences were derived through cycle-sequencing. Sequences were assembled in contigs using the Sequencher ver. 4.1.4 application (Gene Codes) and putative SNPs were identified. A total of 1072 bp from the 1495 bp gene length was represented by high quality sequence. In all 26  
5 SNPs were identified following alignment of sequences from the parental genotypes, at an average incidence of 1 per 40 bp, which is higher than the global frequency of 1 per 60 bp observed through resequencing of 87 genes and c. 76 kb of perennial ryegrass genomic DNA (N.O.I. Cogan, unpublished data). This high SNP frequency, along with the observation of more than two haplotype structures in the NA<sub>6</sub> parent, suggests that  
10 paralogous DNA sequences may have been clustered, due to cross-amplification between gene copies. Polymorphic SNP loci were validated using the single nucleotide primer extension (SNuPe) assay system followed by capillary electrophoresis on the MegaBACE1000 platform (Amersham Biosciences).

Genotypic variation for the SNP locus *LpIRIPNA476* was determined using a sub-set  
15 (96) genotypes of the F<sub>1</sub>(NA<sub>6</sub> x AU<sub>6</sub>) population. The corresponding genomic locus was integrated within the framework of the existing genetic map as previously described (Faville *et al.*, *Theor Appl Genet.* in press). Comparative genomics analysis of *IRIP* genes was performed using the wEST SQL database in the GrainGenes resource (<http://wheat.pw.usda.gov/wEST/>). The nucleotide sequences were used for BLASTN  
20 and TBLASTX analysis in the GrainGenes BLAST page with the search restricted to ESTs that have been assigned to wheat deletion bins Qi *et al* (2003) *Functional and Integrative Genomics* 3:39). The highest matching ESTs were then used to detect the relevant deletion bins using the Mapped Loci query function in wEST SQL.

Genetic map information for the *LpIRIPb* gene was obtained from the second  
25 generation perennial ryegrass genetic mapping population derived from a pair-cross between the genotypes North African<sub>6</sub> (NA<sub>6</sub>) and Aurora<sub>6</sub> (AU<sub>6</sub>). *LpIRIPa* specific locus amplification primers (LAPs) were employed to generate 3 genomic amplicons (Fig. 5A) from each parental genotype. Sequence assembly and analysis revealed 2 distinct *LpIRIP* paralogues *LpIRIP a* and *b*, and a total of 26 SNPs. An assayed SNP locus in  
30 *LpIRIPa* showing an AB x BB segregation structure produced only AB-type F<sub>1</sub> progeny (data not shown), which is diagnostic of multiple gene structure. A total of 8 SNP loci in

*LpIRIPb* showed structures consistent with a single gene copy, but of these 6 showed AA x BB patterns, and could not be genetically mapped in the F<sub>1</sub>(NA<sub>6</sub> x AU<sub>6</sub>) sib-ship. Of the two loci revealing polymorphism within the NA<sub>6</sub> parental genotype (*LpIRIPNA476* and *LpIRIPNA694*) (Fig. 5A), *LpIRIPNA476* was used to genotype the mapping population. The corresponding genomic locus was located on NA<sub>6</sub> LG1 in the terminal location c. 7 cM from locus *xLpesi3f* (Fig. 5B).

#### Expression profiling: RT-PCR?

#### Expression in *E. coli*

The sequence encoding *DalRIpE* and a putative orthologue of histone H3.2 was PCR adapted with the primer pairs  
5' CAGCTTGGATCCATGGCGAACTGCTGTCTGCTA 3' (SEQ ID NO. 137) and  
5' ACTCACAAGCTTAACCTCCTGTCACGACTTTGT 3' (SEQ ID NO. 138); and  
5' AGGAGAGGATCCATGGCGCGTACCAAACAGACC 3' (SEQ ID NO. 139) and  
5' TAATTGAAGCTTTTAGGCGCGTTCGCCACGGAT 3' (SEQ ID NO. 140)  
respectively. They were molecularly cloned into *Bam*HI and *Hind*III restricted pQE-30 (QIAGEN, Germany) and transformed into M15[pREP4]. To induce expression a culture was grown in the presence of ampicillin and kanamycin to mid-log phase, where upon IPTG to 1 mM was added and incubation continued for a further 4 h. Cells were harvested, resuspended in 1 ml extraction buffer and lysed by sonication. The lysate was incubated at 95°C for 10 min., spun at 16,060g for 5 min, the supernatant aliquoted, frozen in liquid N<sub>2</sub>, and stored at -80°C.

## Example 2

## Results

RI activity in *D. antarctica* and *L. perenne* is induced by cold acclimation and present in the apoplasm

- 5 RI assays reveal that *D. antarctica* has activity, induced by cold acclimation, to inhibit further growth of ice crystals following freezing. Given that total leaf extracts from plants grown at 22°C containing 1000  $\mu\text{g mL}^{-1}$  of protein possess no RI activity (Fig. 1) transfer of plants to 5°C for 2 weeks induces RI activity by greater than 64 fold. Furthermore this activity is unaffected by incubation at 95°C for 5 min (Fig. 1). RI activity is also
- 10 induced more than 8 fold in the roots of *D. antarctica* in response to cold acclimation (Table 2). Similarly RI activity in *L. perenne* is below the threshold of detection in the leaves and roots of non-acclimated plants but is induced in excess of 16 and 4 fold respectively following cold acclimation (Table 2).

15 **Table 2.** RI activity<sup>a</sup> in leaves and roots of non-acclimated (grown at 22°C) and cold acclimated (5°C) *D. antarctica* and *L. perenne*.

	Leaves		Roots	
	22°C	5°C	22°C	5°C
<i>D. antarctica</i>	ND <sup>b</sup>	15.6	ND <sup>b</sup>	100
<i>L. perenne</i>	ND <sup>b</sup>	62.5	ND <sup>b</sup>	200

<sup>a</sup> Expressed as lowest concentration of total protein extract ( $\mu\text{g mL}^{-1}$ ) at which activity retained.

<sup>b</sup> No activity detectable at 1000 and 800  $\mu\text{g mL}^{-1}$  for leaves and roots respectively.



Moreover RI activity is present in the extracellular spaces of *D. antarctica* and *L. perenne*. Leaf apoplastic extracts from plants of both species grown at 22°C possess no RI activity whilst activity is induced in response to acclimation at least 73 fold in *D. antarctica* and 1.7 fold in *L. perenne* (Table 3). These correspond to apoplastic protein concentrations, in cold acclimated plants, of 0.31 and 14 µg mL<sup>-1</sup> respectively.

**Table 3.** RI activity<sup>a</sup> in apoplastic extracts of leaves from non-acclimated (grown at 22°C) and cold acclimated (5°C) *D. antarctica* and *L. perenne*.

	22°C	5°C
<i>D. antarctica</i>	ND <sup>b</sup>	89.1
<i>L. perenne</i>	ND <sup>b</sup>	3,830

<sup>a</sup> Expressed as the lowest equivalent wet weight of starting plant material per volume of extract (mg mL<sup>-1</sup>) at which activity retained. <sup>b</sup> No activity detectable at 6,550 and 6,590 mg mL<sup>-1</sup> for *D. antarctica* and *L. perenne* respectively.

Therefore activity to inhibit the consolidation of ice crystals by recrystallisation is induced in response to cold acclimation in both leaves and roots of *D. antarctica*, and to a significantly lesser extent in *L. perenne*. Moreover a significant proportion of this RI activity, particularly in *D. antarctica*, is localised to apoplastic spaces.

**15 IRIP orthologues from *D. antarctica* are predicted to be secreted proteins and contain two types of repeat motif.**

Full length clones of the putative IRIP orthologues *DaIRIPa*, and *DaIRIPd*, e7 and e8 were obtained from *D. antarctica* cDNA and genomic resources respectively. Two genomic clones encoding putative IRIP paralogues *LpIRIP a* and *b* were also obtained from *L. perenne*. In addition many IRIP related sequences have been identified in EST collections from other cereals and grasses. The repeat structures of the longest IRIP

orthologue, *HvIRIP* from *Hordeum vulgare* (barley), and the shortest, *DalRIPd*, are shown in Figure 2A. In all predicted IRIP orthologues (Fig. 2B, C), the C-terminal approximate 120 residues consist entirely of 16 tandem repeats of a degenerate 7-8 amino acid residue motif (the "IRIP repeat") (Fig. 2A). The consensus sequence for the

5 IRIP repeat is SNNTVVSG (SEQ ID NO. 141), with the glycine residue being most conserved (91.9% identity) across all forms. A multiple sequence alignment (Fig. 2B) indicates that relative to IRIP forms in other species IRIPs from *L. perenne* lack the 14th IRIP repeat but have an additional highly degenerate repeat immediately N-terminal to the usual start position of the IRIP domain (Fig. 2B, C).

- 10 Database sequence similarity searches with IRIP sequences reveal that the region N-terminal to the IRIP domain is related to proteins with leucine rich repeat (LRR) motifs. Most closely related is a putative *Oryza sativa* orthologue of a phytohormone receptor, the phytosulfokine receptor (PSKR) (NP\_911036). Regions of significant sequence similarity with IRIPs extend in a discontinuous fashion through the first 17 LRRs of the
- 15 putative PSKR orthologue and approximately 22 residues into a 36 amino acid residue "island" domain (Li and Chory (1997) *Cell* 90:929), where similarity ceases with the advent of the IRIP domain (Fig. 2B, C). With reference to the organisation of LRRs in the *O. sativa* PSKR orthologue IRIPs lack between 8 and 16 of the 17 LRRs (Fig. 2B, C).
- 20 Phylogenetic analysis on the sequences of IRIP orthologues outside of the IRIP domain, together with representative LRR containing proteins reveal that IRIPs fall into a highly robust and distinct clade (Fig. 2D). The sister group to the IRIP clade includes PSKR orthologues and contains LRR receptor-like kinases (LRR-RLKs) of both monocot and dicot origin (Fig. 2D). The most distant clade in this analysis includes
- 25 polygalacturonase-inhibiting protein (PGIP) orthologues, and a LRR containing AFP from *D. carota* (Fig. 2D).

Immediately N-terminal to the LRRs in the predicted IRIPs is a highly conserved 10 amino acid residue motif CCXWEGVXCD (SEQ ID NO. 145) containing 3 invariant cysteine residues (Fig. 2A, B). An additional invariant cysteine residue occurs a further

30 31-32 residues proximal to the N-terminus (Fig. 2A, B). The corresponding cysteine

residues have been shown to form 2 conformationally critical disulphide bridges in the structural determination of the LRR-RLK PGIP of *Phaseolus vulgaris* (Di Matteo et al (2003) *Proc. Natl. Acad. Sci. USA* 100:10124). It is surmised based on their conservation in all IRIP forms that the orthologous residues also participate in structurally important disulphide bonds.

At the N-terminus of all IRIP orthologues is a 20 or 21 amino acid residue region predicted to function as a signal peptide, with a cleavage site between conserved alanine, and threonine or valine residues (Fig. 2B). Consistent with this the mature versions of all full length IRIP forms are predicted to be extracellularly localised. Thus it is likely that IRIPs, are predominantly apoplastic.

#### **Structural modelling of IRIPs predict conformations that are lattice matched to ice surfaces**

Three-dimensional structures of *DalRIPa* and *LplRIPa* (Fig. 3) were constructed by comparative homology modelling. The structural model of *DalRIPa* has three main regions, the double disulphide bonded N-terminal domain, one LRR loop and the IRIP domain (Fig. 3A). *LplRIPa* is similar to the *DalRIPa* in overall structure except for its three LRRs, and additional predicted disulphide bond (Fig. 3B).

The extended  $\beta$ -roll structure of the IRIP domain is predicted to form two surfaces complementary to the prism face of ice, on alternate sides of the domain. Adjacent parallel  $\beta$ -loops are spaced approximately 4.5 Å apart (Fig. 3A, B), whilst threonine and other solvent-accessible residues are arrayed in two ranks on the  $\beta$ -strand faces, spaced 2 residues, or approximately 7.4 Å apart (Fig. 3C). This almost exactly matches the prism ice surface that has repeating structures 4.5 Å parallel to the *a*-axis, and 7.35 Å parallel to the *c*-axis of ice. The putative ice-binding surfaces are stabilised by valine residues that allow tight regular hydrophobic packing of the central core of the  $\beta$ -roll region, and by asparagine residues that participate in hydrogen bonds between adjacent  $\beta$ -strands (Fig. 3D). The highly conserved glycine residues in the IRIP repeat

are structurally important as they form the turns between the upper and lower  $\beta$ -strand faces of the ice-binding  $\beta$ -roll.

Although the conformation of the LRR regions of the *DalRIPa* and *LpIRIPa* models recapitulate the right-handed  $\beta$ -roll of the IRIP repeat they do not engender a flat  $\beta$ -sheet roll. On one side of the  $\beta$ -roll adjacent parallel  $\beta$ -loops are spaced approximately 4.5 Å apart, but on the other, adjacently  $\alpha$ -helical strands cannot pack as closely (Fig. 3A, B). Therefore with each iteration of the LRR the  $\beta$ -roll structure becomes increasingly curved, displaying a concave  $\beta$ -sheet region. As a consequence relative to *DalRIPa* the predicted *LpIRIPa* structure with three LRRs appears to display a less than optimal surface lattice match to the ice prism face (Fig. 3B).

#### Genomic organisation of IRIP related sequences in *D. antarctica* and *L. perenne*

Consistent with the multiple *IRIP* gene variants isolated from them, the genomes of *D. antarctica* and *L. perenne* both appear to harbour multiple IRIP-related sequences. Up to 5 hybridising bands are detected by interrogation of a Southern blot of *D. antarctica* genomic DNA with *DalRIPe* (Fig. 4A). The occurrence of as few as 2 hybridising bands (tracks 3 and 5, Fig. 4A), is evidence that IRIP related sequences may be physically linked on 2 genomic fragments totalling 20kbp. At least 4 *LpIRIPa* sequence related restriction fragments are detected in a sample genome from a heterogeneous breeding population of *L. perenne* (Fig. 4B). The isogenic genomic DNA from a doubled haploid plant, however, exhibits only one strongly hybridising band, with a background of less intense bands (Fig. 4C).

#### SNP-based genetic mapping of *LpIRIPb*

The genetic map position of *LpIRIPb* was determined using single nucleotide polymorphism (SNP). Genetic map information for the *LpIRIPb* gene was obtained from the second generation perennial ryegrass genetic mapping population derived from a pair-cross between the genotypes North African<sub>6</sub> (NA<sub>6</sub>) and Aurora<sub>6</sub> (AU<sub>6</sub>). A total of 8 SNP loci in *LpIRIPb* (Fig. 5A) showed structures consistent with a single gene copy, but

of these 6 showed AA x BB patterns, and could not be genetically mapped in the F<sub>1</sub>(NA<sub>6</sub> x AU<sub>6</sub>) sib-ship. Of the two loci revealing polymorphism within the NA<sub>6</sub> parental genotype (*LpIRIPNA476* and *LpIRIPNA694*) (Fig. 5A), *LpIRIPNA476* was used to genotype the mapping population. The corresponding genomic locus was located on  
 5 NA<sub>6</sub> LG1 in the terminal location c. 7 cM from locus *xLpesi3f* (Fig. 5B).

The DNA sequence from *LpIRIPb* was used to detect putatively orthologous wheat ESTs that had been assigned to the wheat deletion map (Endo and Gill (1996) *Journal of Heredity* 87:95; Qi et al (2003) *Functional and Integrative Genomics* 3:39). The three highest matching ESTs based on TBLASTX analysis were all assigned to deletion bins  
 10 on chromosomes 4AL, 5BL and 5DL (BE48991:  $E = 7 \times 10^{-71}$ ; BF200590:  $E = 5 \times 10^{-58}$ ; BG314423:  $E = 2 \times 10^{-41}$ ). The next lowest matching wheat EST (BG607348:  $E = 2 \times 10^{-40}$ ) detected a deletion bin on chromosome 1BL, as well as 5BL.

#### Expression analysis of *LpIRIPs* and *DalRIPs*

The modulation of steady state levels of *IRIP* gene transcripts in response to  
 15 temperature in *D. antarctica* and *L. perenne*, were investigated. A Northern blot comprising RNA samples extracted from the roots and leaves of *D. antarctica* plants grown at 22°C, 4°C and -16°C probed with full length *DalRIPe7* detects appreciable levels of transcript only in leaves of cold-acclimated plants (Fig. 6A). Quantitative analysis reveal that transfer of plants to 4°C for 2 weeks increases steady state *IRIP*  
 20 mRNA levels approximately 47 fold relative to those grown at 22°C. By contrast *LpIRIPa* transcript levels in *L. perenne* are below the threshold level of detection in leaves, but elevated approximately 4 fold in the roots of cold-acclimated plants relative to those grown at 22°C (Fig. 6B). Thus steady state levels of *IRIP* transcripts are greatly elaborated in leaves of *D. antarctica* in response to cold-acclimation, but only  
 25 moderately so in the roots of *L. perenne*.

#### Heterologously expressed *DalRIPe7* possesses RI activity

Extract from *E. coli* expressing *DalRIPe* was assayed for RI activity. Whilst extracts from cells expressing a putative *D. antarctica* orthologue of histone H3.2 at a

concentration of 400  $\mu\text{g mL}^{-1}$  of protein possess no RI activity, those expressing *DalRIPe7* retain activity down to 100  $\mu\text{g mL}^{-1}$  (Fig. 7). Therefore activity to inhibit further growth of ice crystals following freezing is specifically conferred by *DalRIPe7* and can account, in some part, for this activity *in planta*.

5

### Example 3

#### Discussion

A physiological and functional genomics study in *D. antarctica* has resulted in the identification and characterisation of a gene family encoding IRIPs, the actions of which can account for its tolerance of freezing. *D. antarctica* has activity induced by cold  
10 acclimation, and present in the apoplasm, to inhibit ice recrystallisation thereby minimising the catastrophic plasmolytic consequences of uncontrolled ice crystal growth. This capacity is correlated with the expression of *IRIP* genes, and the primary structure, conformation, localisation, and most significantly, the activity of their products.

#### *IRIP* genes encode proteins with two types of potential ice binding domains

15 We have isolated and characterised putative full length *IRIP* genes from both *D. antarctica* and *L. perenne*. The form reported previously from *L. perenne* (Sidebottom *et al* (2000) *Nature* 406:256), isolated as a protein associated with RI activity, lacked an N-terminal methionine, and was comprised solely of 16 IRIP repeats. The IRIP forms  
20 and LRR domains. Ten LRRs are also the predominant feature of an unrelated (Fig. 2D) ice recrystallisation inhibiting AFP from *D. carota* (Worrall *et al.* (1998) *Science* 282:115; Meyer *et al.* (1999) *FEBS Lett.* 447:171).

Although the LRR domain has the potential to function in ice binding, in the various IRIP forms described here, there is a wide range of variation in its relative contribution to the  
25 overall primary structure. Thus whilst all IRIPs contain 16 IRIP repeats, LRRs are present from as many as 9 iterations in the *H. vulgare* form *HvIRIP* to as few as one in the *D. antarctica* forms, and in the extreme case of *DalRIPd* the residues with similarity

to the LRR number only 17 of the usual complement of 24 or 25 residues (Fig. 2A, B, C).

Despite the apparent plasticity in the number of LRRs, and even their dispensability, other features commonly found in LRR proteins, including the region predicted to participate in 2 disulphide bridges, and the probable signal sequence are invariant in IRIPs, suggesting that they are important for structure/function and/or localisation.

#### **What is evolutionary origin of IRIP genes?**

All plant (and animal) AFPs characterised to date appear to have arisen relatively recently in evolutionary terms by the co-option of existing protein structures (Logson and Doolittle (1997) *Proc Natl Acad Sci U S A.* 94:3485). This is also true of the IRIPs characterised in this study. Outside the IRIP domain itself IRIPs are structurally related to LRR-RLKs, having greatest affinity with orthologues of PSKR, a receptor for the plant hormone phytosulfokine. PSKR first isolated and characterised in *Daucus carota* (carrot) consists of an extracellular domain containing 21 LRRs, a single pass transmembrane domain, and a cytoplasmic serine-threonine kinase domain (Matsubayashi *et al.* (2002) *Science.* 296:1470). Phytosulfokine is a secreted 5 residue sulfated peptide with a key role in cellular de-differentiation and redifferentiation (Matsubayashi and Sakagami (1996) *Proc Natl Acad Sci U S A.* 93:7623). It is difficult to conceive the mechanistic connection between hormonal regulation of cell fate determination and antifreeze activity. A more likely scenario is that a PSKR related protein was co-opted as an AFP, either because of intrinsic structural complementarity to ice crystals, but more credibly as a vehicle to target the "hitch hiking" IRIP domain to the cellular compartment where RI activity is critical, the apoplast. It is possible to envisage an evolutionary scenario whereby a PSKR-like LRR-RLK protein has acquired a novel domain, the IRIP repeat domain, in the process losing its transmembrane and intracellular kinase domains, thereby becoming an untethered apoplastic protein with a novel function. Moreover PSKR-like genes might have been predisposed for such a role because of their expression in organ primordia, cells of which are particularly vulnerable to freezing induced damage. Like the IRIPs, the majority of known plant AFPs are derived from secreted proteins, many being orthologues of pathogenesis-

related proteins (Griffith and Yaish (2004) *Trends Plant Sci.* 9:399), and one, the AFP from *D. carota* (Worrall *et al.* (1998) *Science* 282:115; Meyer *et al.* (1999) *FEBS Lett.* 447:171) also being related to LRR-RLKs.

- IRIPs exhibit plasticity in the number and arrangement of LRRs, more so when compared to their presumed nearest relatives the PSKRs. Thus, there are 17 LRRs in *D. carota* PSKR and its presumed homologue in *O. sativa*, 9 in *HvIRIP*, 3 in the *L. perenne* and *T. aestivum* IRIP versions, and one, or part of one, in the *D. antarctica* forms (Fig. 2B, C). Such evolutionary plasticity in number and arrangement of LRRs has been noted in analyses of LRR-RLKs (see for example Dixon *et al.* (1998) *Plant Cell.* 10:1915)). All the LRRs in extant IRIP forms have high levels of identity and conservation with LRRs in the PSKRs, with no evidence for the addition of any other sequences, LRR or otherwise. Therefore in respect of the principle of maximum parsimony it is most likely that the evolution of IRIPs has resulted from the progressive loss of LRRs rather than their acquisition or rearrangement.
- Because of their evolutionarily recent co-option as AFPs all known plant AFPs exhibit in their sequences clear affinities to particular classes of protein. Although this is true for the PSKR affinities of the LRR related N-terminal part of the IRIPs, the other potential ice binding motif in IRIPs, the IRIP repeat, exhibits no sequence similarity to any reported nucleotide or amino acid sequences. For this reason its origin is unknown.
- Because of its length, the shortest repeat known in an AFP, the IRIP repeat could conceivably be derived from a simple repetitive element of as few as 21 nucleotides in length, such as those found in intergenic DNA. However to date BLASTN searches have failed to identify any closely related genomic sequences.

- Genes encoding IRIP orthologues appear to be monophyletic in origin and confined to the sub-family Pooideae. No sequences related to the IRIP domain have been found in sequence similarity searches of any dicots including *Arabidopsis thaliana*, in the genome of *O. sativa*, or the extensive EST resources derived from *Zea mays* or *Sorghum bicolor*. Furthermore the IRIP clade (Fig. 3B) is highly distinct and deeply rooted supporting the notion that IRIPs arose once early in the evolutionary history of the Pooideae and have subsequently diverged in both copy number and structure. On this



basis *IRIP* genes are predicted to have arisen sometime after the divergence of the Pooideae and Panicoideae 60 mya, but before that of the Triticoideae and Poodae 35 mya (Huang *et al*, (2002) *Plant Mol Biol*. 48:805).

### **Structural modelling predicts that IRIP repeat has greater affinity for ice than LRR**

5 We have used comparative homology modelling to devise a theoretical 3-D structure for full length IRIPs. A truncated version of *LpIRIP* modelled previously (Kuiper *et al* (2001) *Biophys. J.* 81: 3560) did not include the LRR domain, nor the twin disulphide bond-forming N-terminal domain. The structural model demonstrates that both the IRIP and LRR domains can contribute to a common, structurally complementary ice-binding domain.

The predominant ice-binding region is predicted to be the IRIP domain, which presents two ice-binding faces, on either side of the  $\beta$ -roll domain. The putative ice binding surfaces however are not as regular as the stereotypical threonine-X-threonine motifs in beta-roll configurations observed in two unrelated insect AFPs with high TH activity (Graether *et al* (2000) *Nature*. 406:325; Liou *et al*, (2000) *Nature*.406:322).. The *D. antarctica* and *L. perenne* IRIPs exhibit only 30 to 40% threonine at the analogous positions. This is likely due to the differences in the primary function of the proteins. Insect AFPs must provide appreciable TH activity as most insects are not freeze-tolerant. The regularity of the threonine residues on the presenting ice binding surfaces has been implicated in their high TH activity, TH activity having been shown to rapidly decrease with increasing mutational substitution of residues in the ice binding surface (Marshall *et al* (2002) *FEBS Lett.* 529:261). By contrast since *D. antarctica* is freeze tolerant the primary purpose of AFPs in this organism would be to provide RI activity, to avoid the plasmolytic consequences of continued ice crystal growth in already frozen tissue.

In fact, IRIPs may have evolved to have low TH activity as high activity may prove detrimental during the inevitable seasonal freezing of these plants. If a plant were to deploy an IRIP with a relatively high TH activity, the apoplastic fluid of the plant would remain liquid until the temperature dropped below the lower end of the TH gap.

Freezing would then occur much more rapidly than if initiated close to the freezing equilibrium point, and would do so with the spicular dendritic growth observed with other AFPs, potentially doing much mechanical damage to cells.

The LRR domains of full-length IRIPs are also predicted to contribute to ice binding surfaces but not with the inherent structural complementary to the prism face of ice of the IRIP domain. Whilst solved crystal structures of LRRs in proteins form parallel  $\beta$ -sheets on one side of a  $\beta$ -roll, the other side is made up of adjacently packed  $\alpha$ -helical strands (Di Matteo et al (2003) *Proc. Natl. Acad Sci. USA* 100: 10124). As the  $\alpha$ -helical regions cannot pack as closely as the  $\beta$ -sheet regions the  $\beta$  roll structure will curve, displaying a concave  $\beta$ -sheet region, proportional to the numbers of LRRs. Extended curved  $\beta$ -sheet surfaces of LRR regions do not present an optimal surface lattice match to ice, although the AFP from *D. carota* consists predominantly of 10 LRRs. Indeed globular type III fish AFP also does not have an obvious regular ice binding surface and yet displays reasonable TH and RI activity (Baardsnes and Davies, (2002) *Biochim Biophys Acta*. 1601:49).

#### **Genetic mapping of *LpIRIP* and relationship to syntenic cold tolerance and vernalisation QTLs**

In order to determine the location of *IRIP* genes in the genome of *L. perenne* and their proximity to endogenous or syntenic cold tolerance and vernalisation quantitative trait loci (QTLs) *LpIRIPb* was genetically mapped using single nucleotide polymorphism (SNP).

The analysis of SNP variation in *LpIRIP* gene(s) revealed a high level of variation, even compared to average values observed over a large sample of perennial ryegrass. This observation, along with the excess of recovered haplotype structures, is strongly suggestive of multiple gene structure. Although Southern hybridisation analysis indicated a relatively simple genome organisation (Fig. 5C), a number of minor bands were observed, which may correspond to paralogues of the *LpIRIPa* gene. The segregating *LpIRIPNA476* SNP locus may identify such a paralogous sequence, based

on the results of comparative genetics and genomics analysis. Macrosynteny based on heterologous RFLP markers has demonstrated a broad correspondence between each of the perennial ryegrass linkage groups and each of the homoeologous groups of wheat (Jones et al. (2002) *Theoretical and Applied Genetics* 105: 577). On this basis, the SNP locus location on perennial ryegrass LG1 should correspond to a region of conserved synteny with the wheat homoeologous 1S chromosomes. However, comparative genomics has identified wheat sequences assigned to the group 4L and 5L chromosomes, which are related through evolutionary translocations (Devos et al., (1995) *Theoretical and Applied Genetics* 91:282). The location of *LplRIP* ortholoci on Triticeae group 5 chromosomes is also consistent with the detection of QTLs for winter hardiness and frost tolerance on these chromosomes in wheat (Sutka (1994) *Euphytica* 77:277; Galiba et al., (1995). *Theoretical and Applied Genetics* 90:1174; Galiba et al., (1997) *Theoretical and Applied Genetics* 95:265; Toth et al. (2003) *Theor Appl Genet.* 107:509) and barley (Pan et al. (1994) *Theoretical and Applied Genetics* 89: 900; Francia et al. (2004) *Theoretical and Applied Genetics* 108: 670; Reinheimer et al, (2004) *Theoretical and Applied Genetics* 109: 1267), in the same region as the vernalisation response genes that control heading date. Based on conserved synteny, this would predict a location in perennial ryegrass on the upper part of LG4 (Yamada et al. (2004) *Crop Science* 44: 925). The detection of wheat IRIP paralogues on chromosome 1BL, however, indicates the complexity of this gene family in wheat, and suggests that paralogous sequences may be located on other LGs in perennial ryegrass as well. In this interpretation, the non-segregating SNPs may identify variation between LG1 and LG4-located paralogues. In order to test this hypothesis, it would be necessary to identify polymorphic SNPs for the second gene copy in other germplasm. Another possibility is that the LG1-located *xLpiripna476* locus identifies a non-syntenic region. The ends of each LG in perennial ryegrass were enriched for non-syntenic markers, as previously shown for other Poaceae species (Jones et al (2002) *Theoretical and Applied Genetics* 105:577). The closest marker to *xLpiripna476* is an EST-RFLP marker, *xLpesi3f* (Faville et al., *Theor Appl Genet.* In press), which preferentially detected wheat ESTs allocated to deletion bins on chromosome 4A. Finally, it is to be understood that various alterations, modifications and/or additions may be made without departing from the spirit of the present invention as outlined herein.

Documents cited in this specification are for reference purposes only and their inclusion is not acknowledgment that they form part of the common general knowledge in the relevant art.